Globus Online Security Review

Von Welch

February 3, 2012

1 Introduction

This document represents a cybersecurity risk assessment of the Globus Online File Transfer service and associated Website service. It provides a set of concerns related to cybersecurity risk and a set of recommendations to mitigate those concerns and risk. This document also provides an assessment of the cybersecurity-specific documentation provided by the Globus Project for the services and how that documentation improved during the review process.

Please cite as: "Von Welch. Globus Online Security Review. Indiana University ScholarWORKS. February 3, 2012. http://hdl.handle.net/2022/14147"

To check for updates to this document, please see http://papers.vonwelch.com/globus-online-security-review

2 Scope of Review

The review is based on a description, provided by the Globus Project, of the architecture and deployment of the services. The description was provided in the form of written documentation, email communications and a roughly three-hour in-person meeting. This document is written from the point of view of users of the services and operators of data storage services accessed by the services on behalf of their users. It is most accurately described as an architectural security review.

A casual exploration of the services as a user was undertaken, but was not a significant activity in regards to the authoring of the report.

This review is provided as best effort for the benefit of the community. It is provided "as-is" with no warranty expressed or implied. All opinions are those of the author and should not be taken to reflect opinions of any other entity.

2.1 Out of Scope

This review did not, among other things:

- Ensure that the implementation and deployment of the services matched the descriptions given by the Globus Project.
- Evaluate source code for bugs (unsafe coding practices).
- Review the software development, distribution and upgrade process for potential risks.

- Undertake any sort of penetration or other active testing.
- Undertake any review or assessment of software packages, frameworks, operating systems, etc. utilized by the services.
- Consider availability (e.g. resilience to denial of service attacks or non-malicious faults).
- Perform any sort of rigorous cryptographic analysis (e.g. of the web cookie signing).
- Determine if the system satisfies the concerns of any entity (e.g. XSEDE).

3 System Diagram

The following system diagram was created by the author, based on his understanding of the Globus Online system.

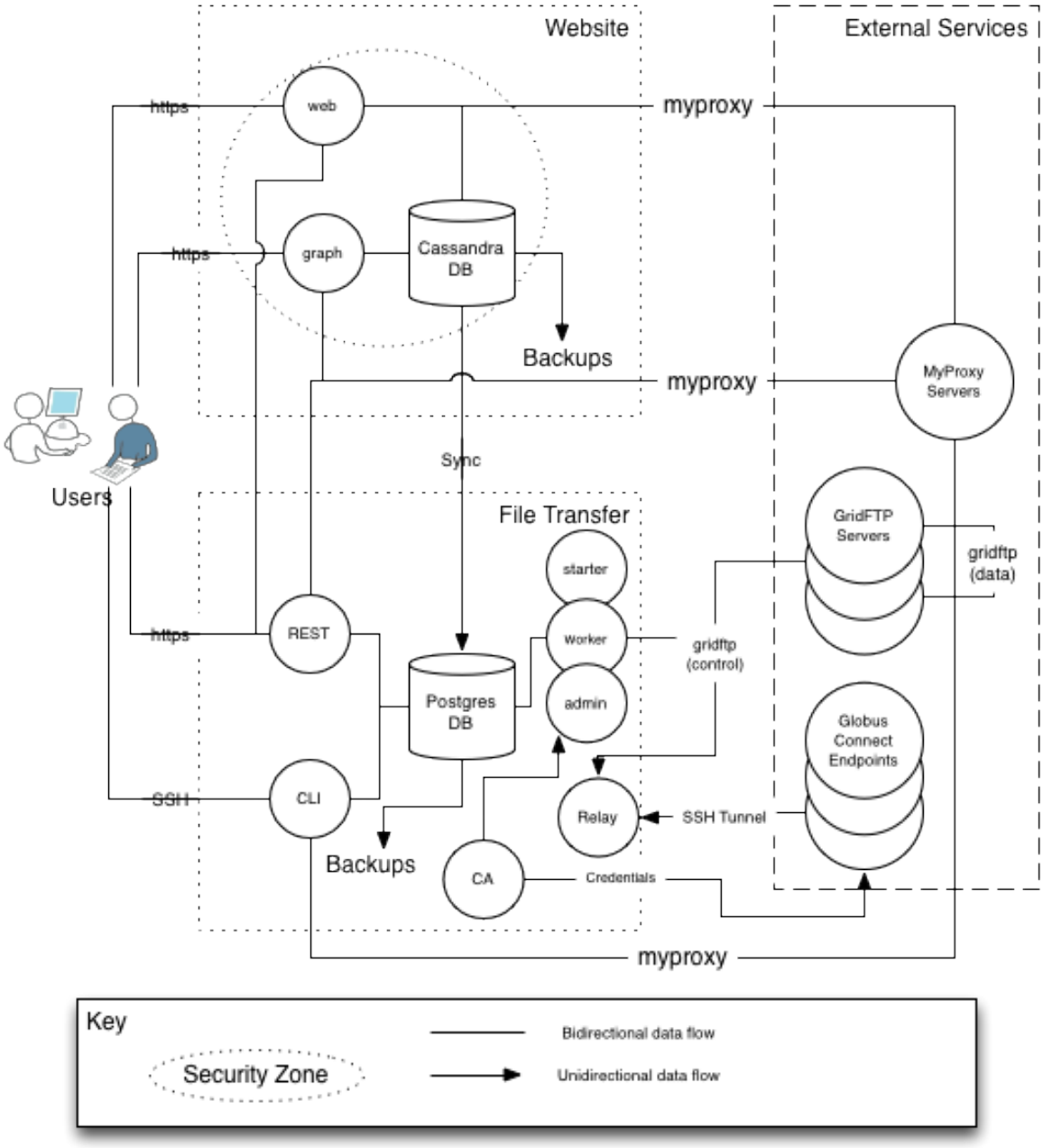

Figure 1: Globus Online File Transfer Service simple architectural view.

4 Overview of the Website

The Website provides two user-facing interfaces:

- Graph: a web server for managing user profiles and identity information (SSH public keys, X.509 identifiers, OpenIds, passwords). User profiles and credentials are stored in the Cassandra database.
- www.globusonline.org: a web server hosting all other web applications for user registration, sign-on, submitting file transfer requests, etc. Requests are submitted to the File Transfer Service via its REST interface.

Other interactions of the Website are:

- Backups of the Cassandra DB are stored in a long-term archive (location unknown).
- MyProxy servers are contacted using information (server name, username, password) provided by users. Returned credentials are stored in memory and may be passed subsequently to the File Transfer Service.
- File Transfer Service REST API. Web applications implement user data movement requests through the File Transfer Service's REST API.
- File Transfer Service Database syncs from Cassandra DB. The File Transfer Service Database maintains a local copy of user profile information held I the Cassandra DB (this is because of historical reasons).

5 Overview of the File Transfer Service

The main functionality of the File Transfer Service (FTS) is provided by a database, which stores information about user requests for data transfers and the status of those transfers, and a set of "Back-end Processing Components" (hence force, "processing components") which act on state in the database.

The FTS provides two user-facing interfaces, both of which exist to allow users to manipulate and query state in the database:

- The Transfer API[1]: A REST API that allows clients to request service. Requests to the REST API are authenticated via X.509 credentials or web cookies generated by the Website.
- The CLI (Command Line Interface)[2]: A SSH-based interface to a restricted shell providing a non-graphical interface for requesting services from the FTS. Connections are authenticated via an SSH key pair or X.509 credential (both managed through the Website).

Other interactions of the File Transfer Service are:

[1] https://transfer.api.globusonline.org/v0.10/doc/
[2] https://www.globusonline.org/usingcli/

- MyProxy Servers: Via the CLI, users can authenticate to MyProxy servers to obtain X.509 credentials for subsequent use to authenticate to GridFTP servers. The credentials are stored in the database for use by the processing components.
- GridFTP Servers: Using user credentials, the processing components contacts GridFTP servers to move data per user direction. These requests are authenticated via X.509 credentials previously obtained via with www.globusonline.org or the CLI.
- Backups of the database are taken hourly. Credentials are not filtered from backups.
- Monitoring: Nagios and Ganglia monitoring services operated by and at Argonne (MCS systems team) monitor the FTS' system health.
- ~~LBNL Monitoring Service: A service operated by LNL regularly, via an SSH connection, pulls data regarding data transfers (name of transferred files,size, speed of transfer, date and time, task identifiers, hostnames and ports).~~ [3]
- Website Cassandra Database: The FTS' database regularly syncs user profile information from the Website Cassandra database.

6 Globus Connect

Users can download and run Globus Connect[4] on computers that are not traditionally GridFTP servers to allow for the transfer of files to and from those computers. Globus Connect can be run in single-user mode, supporting access to files accessible by a given user account, or support multiple users, known as Globus Connect Multi User, which allows access via the accounts of multiple users.

When a user installs Globus Connect, a Globus Online certificate authority (CA) provides a host credential for the deployment. The Globus Online CA certificate is installed as the only trusted CA in the configuration for the deployment. A client distinguished name representing the Globus Online service is placed in the access control list (grid-mapfile) for the installation.

When activated, three things happen: a GridFTP server is launched on the local user computer, Globus Connect establishes an SSH tunnel from the user's computer back to the Globus Online service via the Globus Online Relay, and the GridFTP server on the user's computer is registered with Globus Online as an available endpoint. The purpose of the SSH connection is to allow the access by the File Transfer Service to the Globus Connect GridFTP server through firewalls that may not allow inbound (from the client perspective) GridFTP connections otherwise. With a single-user Globus Connect installation, all these processes run as the user on their computer.

[3] Per discussions with the Globus Project, the documentation is out of date and this monitoring is no longer in place.
[4] https://www.globusonline.org/globus_connect/

When a transfer to or from the Globus Connect computer is initiated, a GridFTP connection is made through the established SSH tunnel. This connection is authenticated via a client certificate created by the Globus Online CA specifically for this purpose. The GridFTP service is expected to authenticate with the credentials created for the installation at the time of its installation by the Globus Online CA. The GridFTP server is configured to only accept incoming connections from the SSH tunnel.

The Globus Connect Multi-User installation is significantly more complicated. Multi-user installations run as root as opposed to a user. The installation includes a MyProxy-based CA, which creates a credential for a local user with a distinguished name containing their local username. This credential is delegated to Globus Online and used to initiate connection back to the Globus Connect GridFTP server. Custom authorization logic extracts the username from the distinguished name and uses that to setuid() to the appropriate user account.

7 Deployment and Administration Summary

The majority of the Globus Online services discussed in this document are deployed on Amazon EC2 and S3 infrastructure. Exceptions are static web content, which is deployed at Argonne, and some monitoring systems, which are also deployed at Argonne.

Administration of the services is done by a combination of MCS Systems staff at Argonne and Globus Project staff at Argonne and U. of Chicago.

8 Existing Globus Online Cybersecurity Documentation

Six documents describing security of the Globus Online Website and File Transfer Services were provided by the Globus Project. The documents appear to be PDFs generated from a wiki page. The documents (and their MD5 hashes) are:

- "Globus Online Security Architecture Overview" dated January 11, 2012. (aa07531ef64f2ef5a2265bd4fae670d0)
- "Website Security". Undated document. (f17213c7c8ae9ba96545b507a68f598c)
- "File Transfer Service Security Overview". Undated document. (0ca6b18caea50127cd6182608ed19a5e) And a revised version of this document entitled "Globus Online File Transfer Service Security Overview". Dated October 7, 2011. (8e1f05138201f3fea9944d4cdd008ad0)
- "Security FAQ". Undated document. (2b599a217e412f23aefb6041927c6952)
- "XSEDE Security Review". Undated Google document printed October 18, 2011 (8bfd865e9d266d35f8f14f9bcda74d85)
- "security overview". Undated document apparently authored by Tom Howe. (61cd9fd78109b0336a5691a3580e3e73)

In this section, a brief summary of each document is provided.

8.1 Globus Online Security Architecture Overview

This Overview document was provided after several review iterations and represented a significant improvement by the Globus project with regards to their documentation. This document contains: (1) a component overview with a Globus Online system diagram and a description of each component, (2) an Identity, Authentication, and Authorization section discussing how each component in turn handles these issues, (3) a Component Interactions and Protocols section describing, for each pair of interacting components, the protocol used and a brief description of the interaction, (4) a Physical Deployment section, a slightly mis-named section discussing not only the deployment strategy but deployment environments, including software stacks, network topologies, monitoring and logging, with ample diagrams, (5) a brief discussion of how Globus Connect works, (6) a description of the Software Engineering process and (7) a Glossary of terms.

A technical improvement described in this document from previous documentation is the addition of OAuth authentication for credential retrieval, mitigating the need for Globus Online to request MyProxy passwords from its users.

8.2 Website Security

The Website Security document is roughly 4 pages long and provides a diagram and description of the Website architecture and deployment. It described the components that make up the Website and their deployment on the Amazon EC2 service. It describes the use of Amazon EC2 Security Groups to provide network-level protection of Website components.

Monitoring, backups and administrative access are also described. The web cookie based authentication is described along with the process for authentication with OpenId and interaction with MyProxy.

8.3 File Transfer Service Security Overview

The File Transfer Service Security Overview (first version) is roughly 4 pages long and provides a description of components making up the File Transfer Service architecture and their deployment on the Amazon EC2 service. A discussion of Amazon EC2 Security Groups is missing and marked as a "TODO". A description of the operating system and software stack used is provided as well as a description of logging. A list of user information assets collected by the service is provided.

The second version of this document primarily added a half-page discussion of actors in the architecture and their privilege levels, presumably in response to a suggestion made by the author of this document.

8.4 Security overview

The rather vaguely named "security overview" document described access control to objects in the database (presumably in the Website). It is a roughly three-page

document with both a data-oriented UML diagram and a deployment diagram. In includes a discussion of password storage and the storage of secret tokens used to authenticate API requests. It also includes a discussion of the storage of temporary credentials and the fact they are encrypted, but no discussion of the key management for that encryption.

8.5 Security FAQ

The Security FAQ is a roughly 3-page document containing 16 questions. The text indicates the questions came from Jim Marsteller in his role as TeraGrid (now XSEDE) security working group lead and from a second individual only identified as "Jim." Of the 16 questions, 10 are completed. The answers are primarily a restatement (and in some cases a direct cut'n'paste) of information found in other documents.

8.6 XSEDE Security Review

The XSEDE Security Review document is a roughly 11-page document. Similar to the Security FAQ, it consists of responses to a request from the XSEDE project. The document is notable in that it contains the only security discussion of the Globus Online Connect service found by this document's author.

9 Assessment of Globus Online Cybersecurity

9.1 Stakeholders

We first consider what entities would be impacted by the security (or flaws there of) of the Globus Online File Transfer Service. The primary stakeholders considered are (numbers are nominal and for identification only):

1. Users of the service who delegate credentials to the service to enable its use.
2. Owners of data accessed by Globus Online.
3. Operators of data services (Globus Connect, GridFTP) accessed by Globus Online.
4. Operators of other services not accessed by Globus Online, but that could be accessed utilizing user credentials held by Globus Online.

9.2 Key Risks

Subjectively, the three main cybersecurity risks, from the perspective of stakeholders, would seem to be (numbers are nominal and for identification only):

1. Maliciously obtaining X.509 credentials (or delegations there of) held by the Globus Online services. (Potentially affects all stakeholders.)
2. Maliciously obtaining user's MyProxy usernames and passwords that are used by the Globus Online services to obtain X.509 credentials on behalf of users. (Potentially affects all stakeholders.)

3. Unauthorized access or modification, via use of Globus Online services in an unintended manner, of user data stored on other services. (Potentially affects stakeholders 1-3.)

9.3 Key Protections

There are a number of key protections in place.

9.3.1 Amazon Security Groups/Firewalls

Amazon security groups act as network firewalls to both limit unneeded access to services running on the Globus Online systems and limit access to services to trusted entities (e.g. database access is limited to trusted clients).

9.3.2 User Authentication

Users are authenticated either directly (e.g. username and password, X.509, OpenId) or subsequently via a web session cookie.

9.3.3 User Interfaces (Web, CLI, REST)

Users interfaces are responsible for limiting a user's actions based on their authenticated identity. This includes establishing what credentials should be used when servicing their requests and what state they are allowed to access.

9.3.4 Encryption of Private Key Material

Private key material is persistent storage is encrypted. In the case of the CA used for the Globus Connect service, the private key must be manually decrypted. User credentials in the database are encrypted with a cryptographic key that is protected with file system permissions.

9.3.5 Use of OAuth to Mitigate use of Passwords

Addressing key threat #2, Globus Online has added support for OAuth[5] so that it does not have to possess user's MyProxy Passwords, reducing their exposure.

9.4 Security Concerns

Given the risks and the system architecture, we turn towards an assessment of security concerns grouped by architectural area. Each concern is labeled Low, Medium or High to indicate the subjective weight given to the concern.

9.4.1 Abuse of User Interfaces (Web, REST, CLI)

Concern #1: The three user interfaces all have the dual responsibility of authenticating the request and ensuring it is legal before passing it on to other components that trust that they have done so. The concern here is a user could craft a malformed request such that the request is undertaken using another user's context and credentials. This indicates this code implementing these interfaces

[5] For more details, see http://www.sciencegatewaysecurity.org/

would be a prime candidate for more extensive architectural and code review. (Medium)

Concern #2: Cross-site request forgery[6] with Web and REST interfaces. After authenticating to Globus Online and delegating a credential, a user could visit another site and be tricked (e.g. through a misleading form or a form submitted automatically with JavaScript) into submitting a request back to Globus Online to move data to an unauthorized receiver, or overwrite data from an authorized source. (Low)

Discussions with the Globus Online developers indicate this concern is currently mitigated by a couple factors:

- The REST API does not accept request as would be generated by a HTML form element; it only accepts AJAX requests. This would prevent a form on another site from submitting a request on behalf of the user.
- The browser same-origin security model should prevent AJAX request from other websites being submitted to Globus Online in a cross-site manner.

Future plans of the Globus Online team include including the session cookie value in the AJAX require to allow for a stronger confirmation that the request was formulated by the Globus Online web site[7]. Requests from external websites will be authenticated with OAuth.

9.4.2 Databases

Concern #3: The databases (Cassandra and the File Transfer database) both hold X.509 credentials and are trusting of their clients to implement appropriate access control. Network security is used to limit access to the databases to those trusted clients. There doesn't appear to be a formed policy on what client components should have what permissions, or an audit/logging process in place. (High)

Concern #4: Password storage is not best practices. Passwords are currently stored as unsalted SHA-512 hashed strings. Best practice would be to use a different salt for each password and a hash algorithm better suited for password storage (e.g. bcrypt[8]). (Medium)

9.4.3 Backups

Concern #5: Backups of the databases are made frequently to help with availability. These backups apparently hold copies of the entirety of database contents (including X.509 credentials). There doesn't seem to be well-defined policy on backup security. (Medium)

---

[6] http://en.wikipedia.org/wiki/Cross-site_request_forgery
[7] https://secure.wikimedia.org/wikipedia/en/wiki/XSRF#Prevention
[8] http://codahale.com/how-to-safely-store-a-password/

9.4.4 Distributed Administration

Concern #6: There are a number of parties involved in the administration of Globus Online and the systems that host the service – the Globus Project at U. Chicago and Argonne, the Argonne MCS Systems Team, and Amazon. This could lead to some increased risk just due to the number of "cooks in the kitchen." For example, who is responsible for coordination in the event of an incident? There does not appear to be an incident response plan in place. (Low)

Concern #7: Using Amazon to host the service introduces the possibility of misbehavior on their part, either organizationally or by a misbehaving employee. There is no evidence that the probability of either of these is anything but a remote concern, but it may be something the Globus Project finds difficult to refute in socializing the security of their service. (Low)

9.4.5 Assessment of Globus Connect

Concern #8: The cybersecurity architecture and trust model of the Globus Connect system is not currently well documented. While the author finds no obvious flaws based on descriptions of the service, it has a complicated architecture, particularly in the Multi-User mode of operation. However, it should be noted that this service only places those who chose to install it at risk. Security documentation for Globus Connect should be drafted and a subsequent evaluation should be undertaken, particularly of its multi-user mode. (Medium)

Concern #9: Globus Connect Multi-user should have appropriate, documented safeguards to ensure that only the local CA is trusted for incoming clients and that what local user accounts can be accessed. Ideally, a limited set of users would be accessible regardless of what username is encoded in the distinguished name and privileged accounts (e.g. root) would be off-limits. This would mitigate the use of Globus Connect Multi-user as a privilege escalation attack. (Medium)

9.4.6 Limiting Authentication Guess[9]

Concern #10: The Globus Web UI does not currently implement protections against brute force password guessing. (Low)

10 Recommendations

10.1 Documentation of Policies and Principles

The Globus Online team should be commended for the security documentation they have produced, in particular the overview document produced after a few iterations of this review. This overview document addresses many of the shortcomings of prior documentation in that it provided a consistent format across multiple components, it provided explicit descriptions of the interactions between

[9] Full credit for the identification of this concern belongs with the PRACE project team. It is included here for completeness.

components, provided numerous diagrams and provided a description of the previously poorly document Globus Connect mechanism, all of which combined to better communicate the architecture.

One shortcoming however is that the current documentation describes practice more than principles and policy. In other words, it captures the "What" but not the "Why". It would be good to start with some high-level principles (e.g. delegated X.509 credentials should only be utilized on behalf of an entity that delegated them; only trusted entities should have communication with the database) and then distill those to polices (X.509 credentials can only be used with jobs whose owner matches the owner of the credentials; only trusted entities should be in the same security group as the database) and then finally to practices (Processing Elements will check the owner of a job and credential and make sure they match; the graphapp security group will only include images X, Y and Z).

10.2 Address Cross Site Request Forgery Issues in User Interface

The project should follow through on their plans to add a session token (and its corresponding validation) to their REST API calls and utilize OAuth to ensure only trusted third party can allow users to submit requests. The project will need to consider what expectations they expect third parties to live up to in terms of ensuring requests are accurate representations of user desire.

10.3 Document a Database-centric Privileges Policy

Securing the database is a key aspect of the Globus Online security architecture. However the database is a trusting entity and any entity that can communicate with it is trusted to enforce at least some portion of the Globus Online security policy. The task of enumerating what privileges each entity should have to the database should be undertaken. To the extent possible, controls should be put into place to actually limit trust in those entities and enforce those limits as close to the database as possible. But even knowing the limits will help with intrusion detection by being able to say when an entity has violated its trust (e.g. an entity that isn't supposed to write does so).

10.4 Firewall/Security Group Testing

Because the database is a trusting entity, Amazon Security Groups are used to implement a network firewall limiting access to it. Monitoring should be expanded to help ensure those firewalls are acting appropriate. In other words, regular testing should be implemented to make help sure there isn't accidental or malicious configuration changes that expose the database to entities who should not have access.

10.5 Logging Strategy

A logging strategy should be created and implemented. The strategy should set forward some well-defined goals, e.g.:

- Be able to trace every action of the File Transfer Service (e.g. data movement) back to a user input.
- Be able to trace every delegation of an X.509 credential to the service and every subsequent use of that credential.
- Be able to trace every access of the database and show it was within the privileges of the client.

Ideally logging would be implemented in such a manner so that is as "write only" as possible, i.e. it would be very difficult for an entity to alter or delete lots no matter how privileged they were. This would probably entail the logging host being located in separate infrastructure.

The EGEE Middleware Security Audit Logging Guidelines[10] may be useful in defining general auditing practices.

10.6 Backup Strategy

There should be a backup strategy that includes whether or not backups contain sensitive data (e.g. X.509 credentials) or such data is filtered out when backups are created. Based on that decision, the backups should be appropriately secured. E.g. Access to backups should be logged.

10.7 Auditing X.509 Credentials

Each credential delegated to Globus Online should be uniquely identifiable by both the public key and the certificate serial number. Logging of all received credentials could help in any investigation since it could serve to help confirm or refute a credential having passed through Globus Online.

10.8 Implement State of the Art Password Storage

Referencing Concern #4, storage of passwords in the database should use best practices.

10.9 Implement Globus Connect Multi-User Safeguards

Referencing Concern #9, Globus Connect Multi-User should have well-documented policies and safeguards to ensure only the appropriate local CA is utilized and privileged local accounts are not accessible.

10.10 Implement Brute Password Guessing Mitigations

Referencing Concern #10, the Globus Web UI should take steps to limit the number of failed authentications to prevent password guessing. Standard practices include locking an account after some number (e.g. 10) failed guesses, or implementing a static or exponential delay between attempts to slow guessing to rates which make it unfeasible.

[10] https://edms.cern.ch/document/793208